

## University of Mississippi eGrove

---

Honors Theses

Honors College (Sally McDonnell Barksdale  
Honors College)

---

2017

# Let's Discuss: Group Size, Course Performance, and Enjoyability in an Interteaching Class

Alexandria Emily Leidt

*University of Mississippi. Sally McDonnell Barksdale Honors College*

Follow this and additional works at: [https://egrove.olemiss.edu/hon\\_thesis](https://egrove.olemiss.edu/hon_thesis)



Part of the [Psychology Commons](#)

---

### Recommended Citation

Leidt, Alexandria Emily, "Let's Discuss: Group Size, Course Performance, and Enjoyability in an Interteaching Class" (2017). *Honors Theses*. 33.

[https://egrove.olemiss.edu/hon\\_thesis/33](https://egrove.olemiss.edu/hon_thesis/33)

This Undergraduate Thesis is brought to you for free and open access by the Honors College (Sally McDonnell Barksdale Honors College) at eGrove. It has been accepted for inclusion in Honors Theses by an authorized administrator of eGrove. For more information, please contact [egrove@olemiss.edu](mailto:egrove@olemiss.edu).

LET'S DISCUSS: GROUP SIZE, COURSE PERFORMANCE, & ENJOYABILITY IN  
AN INTERTEACHING CLASS

By  
Alexandria Emily Leidt

A thesis submitted to the faculty of The University of Mississippi in partial fulfillment of  
the requirements of the Sally McDonnell Barksdale Honors College.

Oxford  
May 2017

Approved by

---

Advisor: Dr. Kate Kellum

---

Reader: Dr. Kelly Wilson

---

Reader: Dr. Michael Allen

© 2017  
Alexandria Leidt  
ALL RIGHTS RESERVED

## ABSTRACT

ALEXANDRIA EMILY LEIDT: Let's Discuss: Group Size, Course Performance, & Enjoyability in an Interteaching Class  
(Under the direction of Dr. Kate Kellum)

In 2013, Saville and colleagues examined whether group size affected course performance in an interteaching based classroom, and found there was no significant difference in course performance between the different sizes of groups. In this replication and extension, we increased the larger interteaching group size from four to six, maintained the small group size at two, and included additional measures of social validity. The students rated their groups each class as a group and individually. Additionally, teaching assistants rated their perception of group effectiveness based on responses to end-of-class questions by individual students. The current study used an alternating treatments design to compared weekly test performance and discussion quality between the small and large interteaching discussion groups. The weekly quiz scores did not show a difference between the two group sizes; however, clear preferences for group size emerged. Findings and implications of the social validity data will be discussed.

## TABLE OF CONTENTS

Introduction.....	1
Method.....	7
Results.....	9
Discussion.....	12
List of References.....	15
Appendices.....	17
Figures.....	21
Tables.....	25

## Introduction

### *Lecturing*

Although faculty use a variety of teaching methods in college classrooms, professors most often tend to use a lecture based classroom (Benjamin, 2002). Lecture remains one of the most frequently used forms of instruction in universities and colleges across the nation and the world (Mowbray & Perry, 2015). Lecture is effective in college classrooms, with research to date suggesting lecture is most effective at the beginning of class. Students show a retention level of 70% within the first ten minutes of a lecture compared to about 20% retention level in the last ten minutes (Lang, 2008). Generally, students in a lecture-based classroom pay their full attention to the professor less than half of the time in a class lasting approximately an hour (as quoted in Lang, 2008; Jones-Wilson, 2005). Lang (2008) suggests that varying the structure of a class will help maintain student attention, and to combine lecture practices with active learning methods would give all students a chance to have their learning needs met. If lecture is most effective in the beginning of classes, and not completely effective as a singular teaching strategy, other teaching methods can be added to increase effective retention of material.

### *Lecture & Active Learning Strategies*

Studies examining classes of a variety of course content show similar results when comparing traditional lecture-based classes to reformed teaching approaches (Franklin, Sayre, & Clark, 2014). In comparing physics students who had been taught in a traditional lecture-based classroom to those taught in a classroom that actively engaged the students through class participation, students showed a similar comprehension in both environments (Franklin, Sayre, & Clark, 2014). However, students in the active-

engagement classroom showed greater retention of the material long term than their peers in the traditional lecture-based classroom (Franklin, Sayre, & Clark, 2014).

Lecture is likely useful for facilitating student comprehension, but retention beyond the term in which the class is taken is often the goal of enrolling in a class. Universities and colleges are more than facilities for learning; they are facilities for students to learn how to use knowledge as a tool to benefit society as a whole (Lagemann & Lewis, 2012). Many students attend higher education in order to reap the individual benefits of a good job or financial stability, and these benefits are advantaged when students learn and apply material effectively (Lagemann & Lewis, 2012).

#### *Difficulties with Non-Traditional Teaching Methods*

If the purpose of secondary and higher education is to learn competency in a given field, then professors may wish to consider the data concerning human learning and behavior where there is support for more effective ways to educate students (Saville, Lambert, & Robertson, 2011). These data are primarily produced by those involved in the field of psychology, and professors involved in this field could carve a path for other disciplines through demonstration of more effective teaching methods. However, it is not enough for those who produce evidence in support of non-traditional teaching methods to disseminate their findings. Professors of all disciplines can be encouraged by their institutions and fellow colleagues to seek out readily available information concerning non-traditional teaching methods.

Saville et al. (2011) highlight five key reasons as to why teachers and students may not look to more effective teaching practices. First, the traditional structure of the academic calendar may not match the time needed to properly instruct through non-

traditional teaching methods (Saville et al., 2011). Second, lecture classrooms are the standard when looking at the structure of education. It is not easy for professors to change how they were instructed and their preferred system of teaching (Saville et al., 2011). Third, the resistance to deviate from the lecture-based classroom is not entirely the choice of the professor. As students are the main concern of the class, the professor may not want to change what is accepted by students, regardless of whether learning is improving (Saville et al., 2011). Fourth, the reason some professors decide to teach is to impart knowledge to students. Professors are experts in their field, and it is not easy for an academic who entered a field with the intent of being an expert to want to surrender control of the classroom to students who do not possess the same expert field knowledge (Saville et al., 2011). Finally, if the professor is not educated properly on the functions of the behavioral practices, misinformation can dissuade a professor from exploring the benefits of intermingling behavioral practices into the academic curriculum (Saville et al., 2011). Implementation of non-traditional instructional designs in classrooms is difficult, but not impossible. Through behaviorally backed practices that are flexible, professors can enjoy the benefits of more effective teaching.

### *Interteaching*

Interteaching is an instructional style based on the principles of behavior analysis with the purpose to facilitate discussion between students in order to improve material comprehension (Boyce & Hinline, 2002). This discussion is referred to as an interteach, a conversation between students used to discuss and encourage application of concepts covered in articles or a class textbook (Boyce & Hinline, 2002). From these discussions, students are able to understand what material is unclear and needs further explanation by



the professor, and what material is comprehended. This understanding allows for lectures to be beneficial at supplementing the material that is not clear (Boyce & Hineline, 2002).

Interteaching has shown better performance outcomes when compared to lecture and/or simply reading the material (Saville et al., 2005). While interteaching is shown through research to be an effective alternative to the lecture-based classroom, the implementation of interteaching has been slow. Education as an administrative structure has not experienced significant changes in basic practices despite the persistence of problems. A main issue in the education system identified in the 1950s by B.F. Skinner highlights the lack of positive reinforcement used in classrooms (Saville et al., 2011). Several non-traditional teaching methods were created out of Skinner's criticism, such as precision teaching and personalized system of instruction, but these methods were not the finite solution to the problems in traditional education practices (Saville et al., 2011). As a response to educators voicing their dislike of previous non-traditional methods, Boyce & Hineline (2002) took into consideration the best characteristics of the methods created from Skinner's criticism, and attempted to market interteaching as the new method more suited to the needs of the traditional educator.

#### *Interteaching Classroom Procedure*

Interteaching is a combination of lecture-based teaching practices and active student engagement (Brown, Killingsworth, & Alavosius, 2014). Before class begins, students are provided discussion guides that contain questions and material relevant to the reading. The purpose of the discussion guide is to assess knowledge regarding terms and concepts, and provide questions on how to properly apply the terms and concepts (Saville et al., 2005). A short lecture is given in the beginning of class to review concepts covered

in the previous class, and clarify student confusion, if any. Following the review lecture, students break out into discussion groups focused on the discussion guide questions (Saville et al., 2005). Discussion groups are an important part of the class and allow the students to gain more knowledge through questions regarding application of concepts rather than definitions of terms (Boyce & Hineline, 2002). Saville et al. (2005) suggests that discussion between learning groups should comprise 75% of class time depending on the class length and amount of material needing review in the clarifying lecture. Due to the amount of preparation and involvement required of the professor, teaching assistants are often recruited to assist in the daily class activities, and typically assist the instructor with the facilitation of discussion or clarifying concepts while students are in their discussion groups (Saville et al., 2005). Once the concepts and topics of the discussion guide have been reviewed in groups, students fill out an interteaching record, which allows the instructor to assess which concepts should be included in the next clarifying lecture (Saville et al., 2005).

Even though interteaching has been shown to improve retention and classroom performance in students, it is not a widely-supported practice by faculty in colleges and universities (Brown, Killingsworth, & Alavosius, 2014). The time commitment required to run an interteaching class is significant, and the professor has to be willing to adapt to the class as it progresses (Brown, Killingsworth, & Alavosius, 2014). Each class requires a discussion guide be created with questions to assess and teach students, and the interteaching record must be collected and reviewed by the professor in order to properly understand what concepts need to be reviewed next class in the clarifying lecture. A critical component to the interteaching model is the group discussion following the

professor's lecture. Although already a successful teaching method in many college classrooms, interteaching is subject to improvements. Discussion is an important part of the interteaching structure as it gives students the ability to collaborate and learn through peer interaction.

### *Current Study*

Depending on the preference of the professor, the number of students in each discussion group fluctuates. A study published in 2013 examined whether there would be a statistically significant difference in exam scores when students were placed in groups of two or four students (Truelove, Saville, & Van Patten, 2013). No significant difference was found between the groups, and this is the basis for the current study.

The purpose of this study was to examine the effects of group size on course performance and the enjoyability of the class. Truelove and colleagues (2013) examined the differences in test scores between groups of two people and groups of four people. In the current study, the small group was maintained at two people and the large group was expanded to six people. We examined whether there was a relationship between discussion group size and course performance as Truelove and colleagues (2013) did with modifications to the large group size. Additionally, we examined the relationship between discussion group size and the quality of discussion as reported by the group, the individual student, and the teaching assistant.

## Method

### *Participants*

Sixty undergraduate students enrolled in a Psychology of Learning class participated in this study. The investigation was classified as “exempt” from the Institutional Review Board (IRB) because it was conducted in a classroom and was considered educational practice. Students were provided an information sheet that outlined the purpose and procedure of the study, and instructions about how to have individual data not included in the final analysis. No student asked for his or her data to be removed from analysis. Students ranged in classification from sophomore to senior. There were nine teaching assistants who enrolled for internship credit and assisted in data collection each class period.

### *Materials & Design*

We used an alternating treatments design to examine the relationship between discussion group size and course performance as well as enjoyability in an interteaching class. We used three social validity measures and weekly quiz grades that served as our course performance measures. The first social validity measure used was a group rating sheet, which asked the group to rate the quality of group discussion. The rating scale ranged from 1-7 (1= poor quality of discussion, 7 = superb quality of discussion) (See Appendix A). The second social validity measure was an individual rating sheet that asked the participant to rate the quality of group discussion independent of their group members. The individual rating scale ranged from 1-10 (1= poor quality of discussion, 10 = superb quality of discussion) (See Appendix B). The final social validity measure used was a rating scale completed by each teaching assistant. Each group the teaching assistant

talked to at the end of class question was rated on how well each member answered an end-of-class question. The rating was on a scale of 1-10 (1= poor quality of discussion, 10 = superb quality of discussion) (See Appendix C).

### *Procedure*

When students enter the class, their discussion guides are examined for completion followed by a short clarifying lecture to review past material and introduce new concepts covered in the most recent reading and discussion guide. Depending on the condition, students were divided into groups of two or six people. In weeks when students were organized into two-person groups, students were able to select their partner. In weeks when students were organized into groups of six, notecards were used to divide the students into groups, and the students did not have a choice with whom they were grouped. At the end of each group discussion, three social validity measures were collected. First, each group completed a record sheet that included a group rating measure. The record sheet was given to the teaching assistant at the conclusion of the group discussion. Once the Teaching Assistant collected the rating sheet, the Teaching Assistant rated the quality of answers given by each group member. The questions asked by the Teaching Assistant were based on the day's discussion guide. While the Teaching Assistant rated the quality of the answers from the group, each student individually rated their perception of the quality of group discussion. All students completed a group rating and individual rating at the end of each class discussion, and course performance was measured using scores from the end of week quizzes.

## Results

### *Weekly Quizzes*

Weekly quiz data analyzed through visual analysis showed no apparent difference between the weeks when students were in groups of two compared to when students were in groups of six. All quizzes were scored out of fifteen points with the exception of Week 08. The data for Week 08 were adjusted so that they reflected the scores to be out of fifteen instead of twenty-five. At the conclusion of the group size intervention, students were allowed to choose the number of people in their groups for the remaining three weeks of the semester.

Week 01 placed students in groups of two people. The quiz data showed a mean score of 14.26 out of 15 points with a standard deviation of 0.85. Week 02 placed students in groups of two people again, and showed a mean score of 10.57 out of 15 points with a standard deviation of 3.79. Students were placed in groups of six in Week 03. The mean quiz score was 12.36 with a standard deviation of 2.12. During Week 04, students were placed in groups of two. The mean quiz score for the week was 10.77 with a standard deviation of 3.69. Week 05 placed students in groups of six people. The mean quiz score was 12.71 with a standard deviation of 2.39. The next week, Week 06, students returned to groups of two people. The mean quiz score was 13.28 with a standard deviation of 1.83. Week 07 placed students in groups of six. The mean quiz score was 13.80 with a standard deviation of 1.83. Weekly Quiz 08 was out of twenty-five points originally and was adjusted to reflect the quiz score being out of fifteen points to match the rest of the quiz data that was collected. Students were placed in groups of

two students and the adjusted mean quiz score was 12.06 with a standard deviation of 5.63. Week 09 placed students in groups of two. The quiz scores for the week showed a mean of 11.53 with a standard deviation of 2.03. Week 10 was the final week students were placed in groups of six people. The mean quiz score was 11.35 with a standard deviation of 2.31. Finally, Week 11 placed students in groups of two. The mean score on the weekly quiz was 9.38 with a standard deviation of 3.05. For Weeks 12-14, students were allowed to choose how many people were involved in their class discussions. The mean quiz score for Week 12 was 10.89 with a standard deviation of 3.51. The mean quiz score for Week 13 was 12.32 with a standard deviation of 2.33. The mean quiz score for Week 14 was 10.81 with a standard deviation of 2.98 (See Table 1).

#### *Teaching Assistant Rating*

Visual inspection of rating data collected from the Teaching Assistants did not show apparent differences between two-person groups and six-person groups. The mean quality of discussion rating of the Teaching Assistants was 7.6 for both groups of two and groups of six on a scale of 1-10 (1 = poor quality of discussion, 10 =) consistently over the 11-week intervention. There are some outlying ratings where the Teaching Assistants rated the groups lower on quality of discussion, however the ratings overall do not show a reliable deviation from the mean rating.

#### *Group Rating*

The group rating data did not show apparent differences in preference between the two-person groups and six-person groups. To analyze the group rating data, boxplots superimposed over violin plots were created. The data show a mean rating of 6.2 on a

scale of 1-7 for groups of two and a mean rating of 6.1 on a scale of 1-7 for groups of six (1= poor quality of discussion, 7 = superb quality of discussion).

### *Individual Rating*

The individual rating data shows some differences between groups of two and groups of six. Boxplots overlaid on top of violin plots were created to analyze the data. With the exception of weeks ten and eleven, the data suggest that students prefer the smaller two-person groups compared to the six-person groups. When students were in groups of two people, their individual perception of the quality of discussion showed a mean of 8.5 on a scale of 1-10 (1 = poor quality of discussion, 10 = superb quality of discussion). When students were placed in groups of six, the mean rating was 7.4 on a scale of 1-10.



## Discussion

This study investigated whether there is a relationship between discussion group size and the quality of group discussion. The data show that through self-report on quality of discussion, students indicated preference of two-person groups through the individual rating data. The self-report data collected from the group record sheet did not show apparent differences in how students rated quality of discussion as a group between small and large group sizes. The ratings collected from the Teaching Assistants did not show different student preferences for small and large groups. Weekly quiz scores collected at the end of each week did not show differences between the two and six-person groups.

### *Rationale and Limitations*

As with any investigation, there were uncontrolled variables that likely influenced the results. For example, it was likely disrupting for students to move seats in the auditorium classroom in order to form groups. The seating in the classroom was fixed and required students to move themselves instead of their selected seat. In weeks where students were in two-person groups, students were allowed to choose their partner for the day. When students were placed in six-person groups, they were assigned their groups by the color notecard they received on that day. This may have brought discomfort as students had to move around much more during the classes organized in six-person groups, and could have affected the individual ratings given by each student at the end of each day. Students were vocal about their dislike of the weeks when they were placed in large groups and the small two-person groups were communicated to be favorable through the comments left on the individual rating sheets.

Additionally, students are more easily able to use discussion time to talk about topics other than those listed on the discussion guide while in groups of two. There is only one other person to account for, and it is likely that students in two person groups were able to easily agree on what they would rate the quality of discussion on their individual sheets, as well as the group rating sheet. While in six person groups, students were not able to stray from the topic without another group member noting it in their personal rating sheet.

Finally, the individual and Teaching Assistant rating measures were implemented in the third week of the investigation. This may have disrupted the flow established by students in the first two weeks of class. The students experienced two weeks of class where they worked in groups of two and this became the routine for them. When the larger group intervention was implemented, the aversive function of being in large discussion groups was perhaps offset by the appetitive function of being in small discussion groups. One final limitation was that the structure of the class follows an assignment menu where students can choose assignments they would like to complete rather than a list of assignments that require completion. Students are not required to complete the weekly quizzes, and group size may have influenced the number of students who took the weekly quiz from week to week. We are unsure if extraneous variables affected the self-report and weekly quiz data collected at the end of the investigation.

### *Implications*

The data did not show apparent differences between groups for the mean weekly quiz scores. However, the social validity measures completed by each individual student following group discussion show student preference for small groups as opposed to large

groups. In addition to the numeric rating, students often commented on their rating sheets to communicate their dislike for the large groups, and asked to be placed back in small groups for discussion. In some cases, students reported they did not enjoy being in larger groups because all members did not contribute equally to the group discussion.

In addition to group size, the ability for students to pick their partners as opposed to having their partners assigned might have an effect on the quality of group discussion, regardless of whether the students are in large or small group sizes. In this investigation, students were allowed to pick their partner when they were organized in two-person groups, but were assigned their partners when the class was organized in six-person groups. Additionally, it would be interesting to examine whether the instructional design of interteaching in a non-psychology based classroom yields similar results. The use of interteaching as an instructional method across disciplines in colleges and universities would benefit students in the amount of material retained over time.

## List of References

- Benjamin, L. T. (2002). Lecturing . In S. F. Davis, & W. Buskist, *The Teaching of Psychology: Essays in Honor of Wilbert J. McKeachie and Charles L. Brewer* (pp. 57-67). Mahwah, New Jersey : Lawrence Erlbaum Associates, Inc.
- Bok, D. (2013). Undergraduate Education. In D. Bok, *Higher Education in America* (pp. 75-80). Princeton University Press .
- Boyce, T. E., & Himeline, P. N. (2002). Interteaching: A Strategy for Enhancing the User-Friendliness of Behavioral Arrangements in the College Classroom. *The Behavior Analyst* , 215-226.
- Brown, T. W., Killingsworth, K., & Alavosius, M. P. (2014). Interteaching: An Evidence-Based Approach to Instruction. *International Journal of Teaching and Learning in Higher Education*, 26(1), 132-139.
- Franklin, S. V., Sayre, E. C., & Clark, J. W. (2014). Traditionally taught student learn; actively engaged students remember. *American Association of Physics Teachers*, 82(8), 798-801.
- Guy, R., & Marquis, G. (2016). The Flipped Classroom: A Coparison of Student Performance Using Instructional Videos and Podcasts versus the Lecture-Based Model of Instruction. *Issues Informing Science and Information Technology*, 13, 1-13.
- Lagemann, E. C., & Lewis, H. (2012). What is College For? The Public Purpose of Higher Education. In E. C. Lagemann, & H. Lewis, *renewing the Civic Mission of American Higher Education*. New York, New York: Teacher's College Press.

- Lang, J. M. (2008). *On Course: A Week-By-Week Guide to your First Semester of College Teaching*. Cambridge, Massachusetts, United States of America: Harvard University Press.
- Mowbray, R., & Perry, L. B. (2015). Improving lecture quality through training in public speaking. *Innovations in Education and Teaching International* , 207-217.
- Querol, B. I., Rosales, R., & Soldner, J. L. (2015). A Comprehensive Review of Interteaching and Its Impact on Student Learning and Satisfaction. *Scholarship of Teaching and Learning in Psychology*, 1(4), 390-411.
- Saville, B. K., Zinn, T. E., & Elliot, M. P. (2005). Interteaching Versus Traditional Methods of Instruction: A Preliminary Analysis. *Teaching Psychology*, 32(3), 161-163.
- Saville, B. K., Lambert, T., & Robertson, S. (2011). Interteaching: Bringing Behavioral Education Into The 21st Century. *The Psychological Record*, 61, 153-166.
- Truelove, J. C., Saville, B. K., & Van Patten, R. (2013, May). Interteaching: Discussion group size and course performance. *Journal of the Scholarship of Teaching and Learning*, 13(2), 23-30.

## Appendices

Appendix A  
Group Rating Sheet

**PAIR DISCUSSION - DAILY RECORD SHEET**

Day # \_\_\_\_\_ Date of discussion \_\_\_\_\_

Participants \_\_\_\_\_  
\_\_\_\_\_

For TA use only:

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

Duration of discussion \_\_\_\_\_

Sufficient time provided? Yes No

Quality of pair discussion (circle one) 1 2 3 4 5 6 7  
poor OK  
superb

If “poor” or “superb,” what contributed to the quality?

List the **top 3** questions (if any) that you would most like reviewed in the next lecture

What do you think the “main point” of today’s reading and prep guide was?

What is something interesting you learned in class today (you must list something)

List at least one reason why you are glad you came to Learning today

Other comments and/or suggestions

Appendix B  
*Individual Rating Sheet*

Quality of discussion:

1	2	3	4	5	6	7	8	9	10
poor				OK					superb



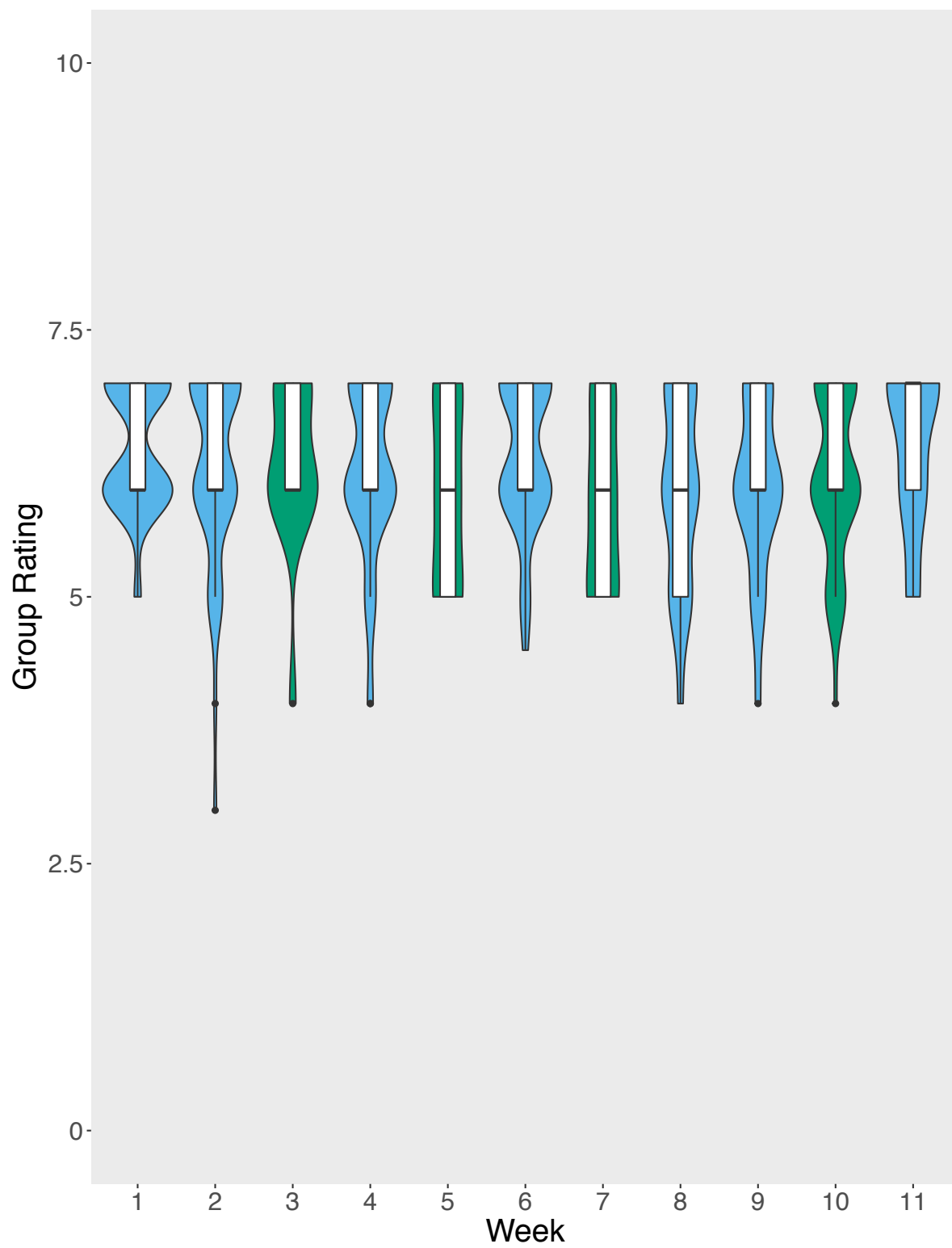
Appendix C  
*Teaching Assistant Rating Sheet*

What questions did groups ask about?

Group Size:  
Quality of Discussion

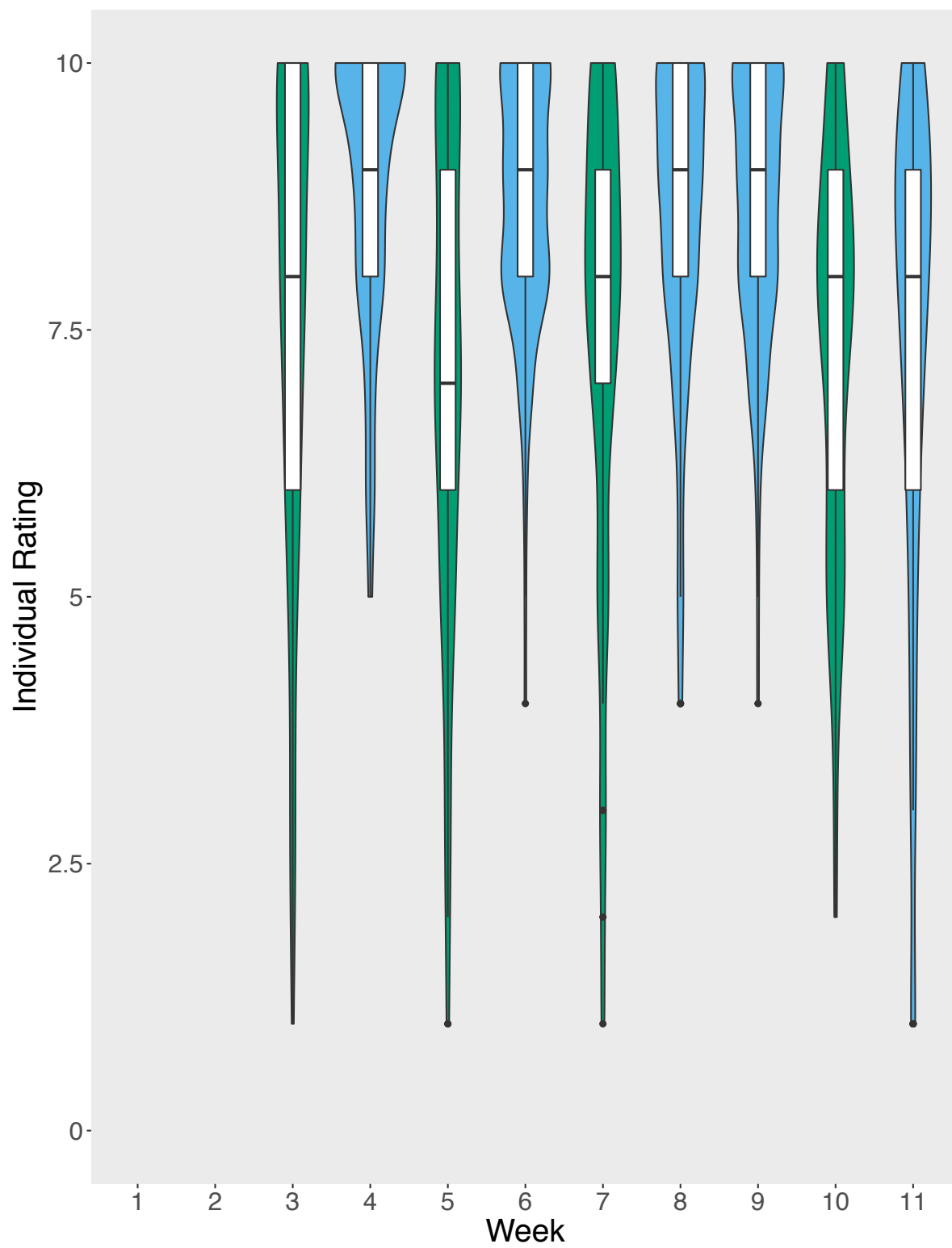
1	2	3	4	5	6	7	8	9	10
poor				OK					superb

Figure 1



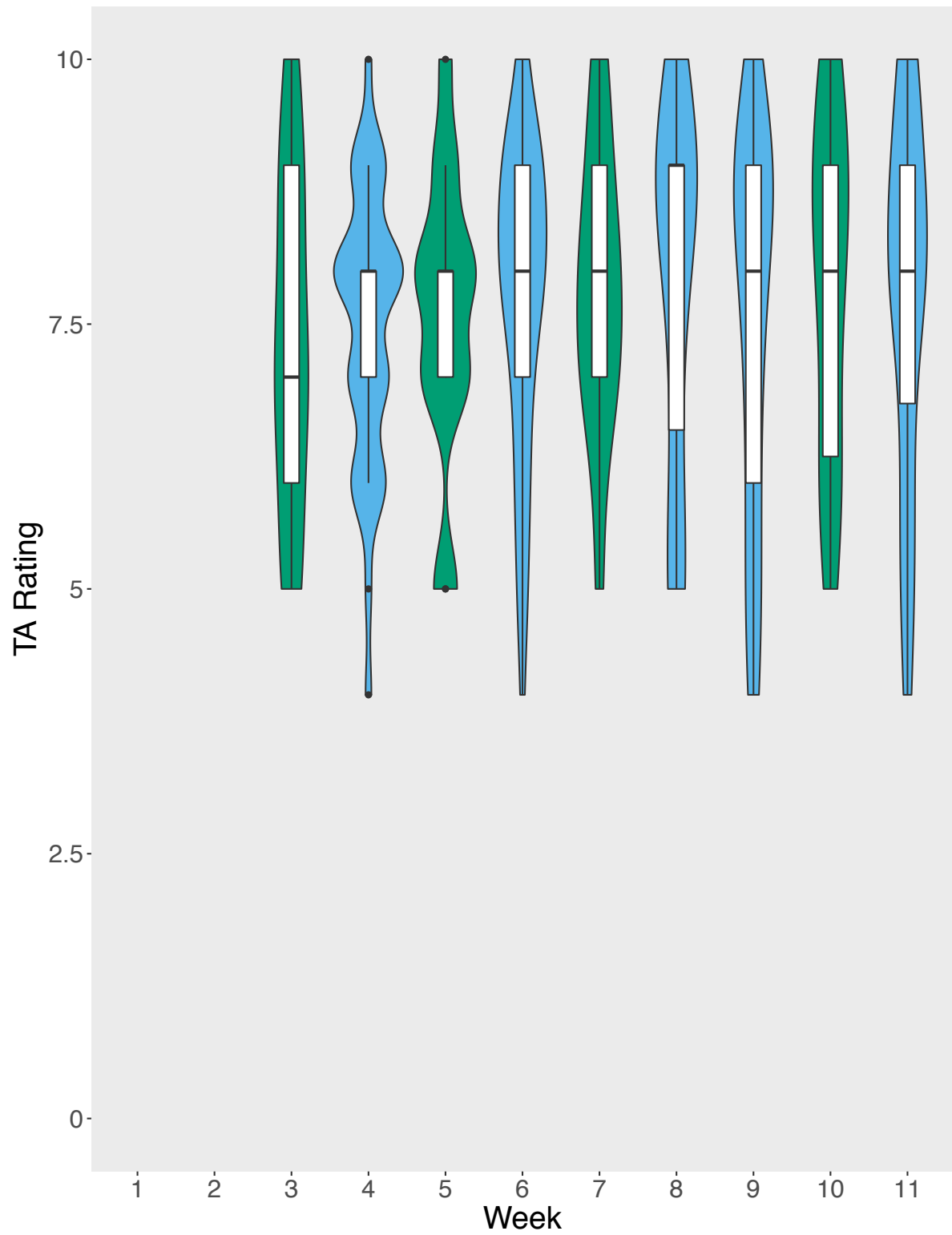
Mean Group Quality of Discussion Rating

Figure 2



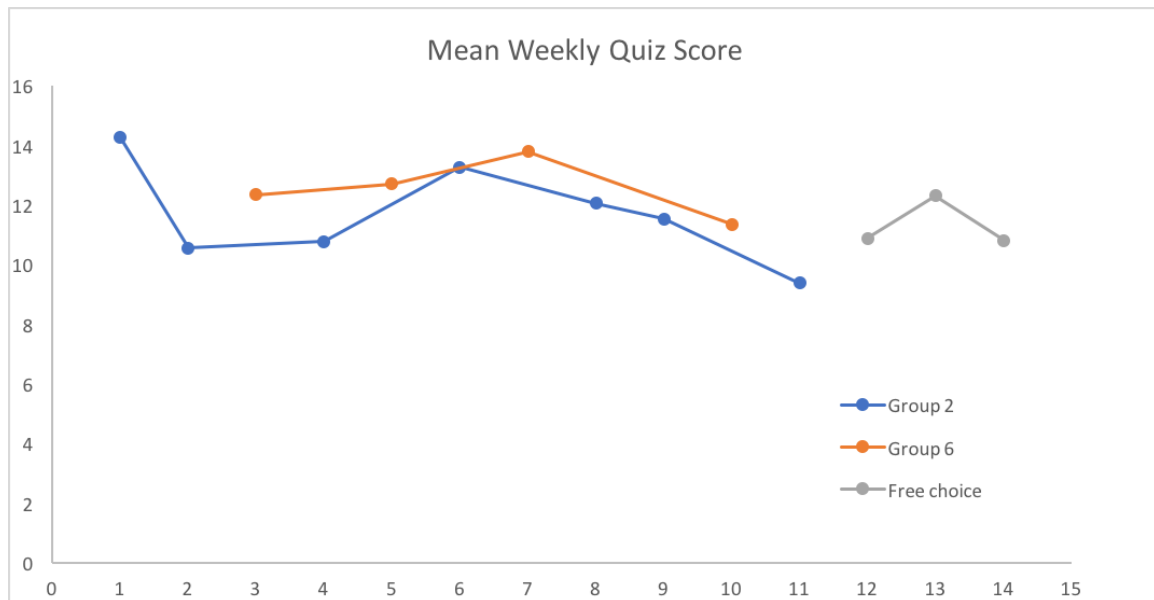
Mean Individual Quality of Discussion Rating

Figure 3



Mean Teaching Assistant Quality of Student Answers Rating

Figure 4



Graph of Mean Weekly Quiz Score

Table 1

	Group Size 2	Group Size 6	Free Choice
Week 1	14.26		
Week 2	10.57		
Week 3		12.36	
Week 4	10.77		
Week 5		12.71	
Week 6	13.28		
Week 7		13.80	
Week 8	12.06		
Week 9	11.53		
Week 10		11.35	
Week 11	9.38		
Week 12			10.89
Week 13			12.32
Week 14			10.81

Mean Weekly Quiz Score